



HAL
open science

Interactive Audio Augmented Reality in Participatory Performance

Anna N Nagele, Valentin Bauer, Patrick G T Healey, Joshua D Reiss, Henry Cooke, Tim Cowlshaw, Chris Baume, Chris Pike

► **To cite this version:**

Anna N Nagele, Valentin Bauer, Patrick G T Healey, Joshua D Reiss, Henry Cooke, et al.. Interactive Audio Augmented Reality in Participatory Performance. *Frontiers in Virtual Reality*, 2021, 1 (610320), 10.3389/frvir.2020.610320 . hal-04447862

HAL Id: hal-04447862

<https://universite-paris-saclay.hal.science/hal-04447862>

Submitted on 8 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Interactive Audio Augmented Reality in Participatory Performance

Anna N. Nagele^{1*}, Valentin Bauer^{2*}, Patrick G. T. Healey¹, Joshua D. Reiss¹, Henry Cooke³, Tim Cowlshaw³, Chris Baume³ and Chris Pike³

¹Media and Arts Technology, Electronic Engineering and Computer Science, Queen Mary University of London, London, United Kingdom, ²LISN, CNRS, Université Paris-Saclay, Orsay, France, ³BBC Research and Development, London, United Kingdom

Interactive Audio Augmented Reality (AAR) facilitates collaborative storytelling and human interaction in participatory performance. Spatial audio enhances the auditory environment and supports real-time control of media content and the experience. Nevertheless, AAR applied to interactive performance practices remains under-explored. This study examines how audio human-computer interaction can prompt and support actions, and how AAR can contribute to developing new kinds of interactions in participatory performance. This study investigates an AAR participatory performance based on the theater and performance practice by theater maker Augusto Boal. It draws from aspects of multi-player audio-only games and interactive storytelling. A user experience study of the performance shows that people are engaged with interactive content and interact and navigate within the spatial audio content using their whole body. Asymmetric audio cues, playing distinctive content for each participant, prompt verbal and non-verbal communication. The performative aspect was well-received and participants took on roles and responsibilities within their group during the experience.

Keywords: audio augmented reality, 3D audio, audio-only game, experience design, immersive theater, interaction design, participatory performance, interactive storytelling

1 INTRODUCTION

The emergence of affordable consumer technologies for interactive 3D audio listening and recording has facilitated the delivery of Audio Augmented Reality (AAR) experiences. AAR consists of extending the real auditory environment with virtual audio entities (Härmä et al., 2004). The technology has been applied to fields such as teleconferencing, accessible audio systems, location-based and audio-only games, collaboration tools, education or entertainment. The main difference of AAR toward traditional headphones is that it does not create a sound barrier between private and public auditory spheres. AAR technology should be acoustically transparent (Härmä et al., 2004), i.e., the listener should perceive the environment as if they had no headphones, without being able to determine which sound sources are real and which ones are not. This feature remains one of the main technical challenges in the field (Engel and Picinali, 2017).

Research in the field of AAR has focused on technical features and interactive possibilities with devices and media, with little exploration of the user experience (Mariette, 2013). It has explored perceived sound quality (Tikander, 2009), discrimination between real and virtual sounds, perception of realism (Härmä et al., 2004) or user adaptation to wearable AAR (Engel and Picinali, 2017). Few interactive story-led experiences have been studied (Geronazzo et al., 2019). While many headphone experiences remain individual (Drewes et al., 2000; Sawhney and Schmandt,

OPEN ACCESS

Edited by:

Victoria Interrante,
University of Minnesota Twin Cities,
United States

Reviewed by:

Tapio Lokki,
Aalto University, Finland
Gershon Dublon,
Massachusetts Institute of
Technology, United States

*Correspondence:

Anna N. Nagele
a.n.nagele@qmul.ac.uk
Valentin Bauer
valentin.bauer@limsi.fr

Specialty section:

This article was submitted to
Virtual Reality and Human Behavior,
a section of the journal
Frontiers in Virtual Reality

Received: 25 September 2020

Accepted: 23 December 2020

Published: 12 February 2021

Citation:

Nagele AN, Bauer V, Healey PGT,
Reiss JD, Cooke H, Cowlshaw T,
Baume C and Pike C (2021) Interactive
Audio Augmented Reality in
Participatory Performance.
Front. Virtual Real. 1:610320.
doi: 10.3389/frvir.2020.610320

TABLE 1 | Overview of interactive audio technologies used in the context of theater and performance.

Production	Audience experience	AAR	Interaction techniques	Hardware	Tracking	Location
ANNA https://www.nationaltheatre.org.uk/shows/anna	Seated audience is wearing headphones to hear the play from the protagonist's point of ears.	Binaural audio recorded in real time from the protagonist's point of ears.	None	Wearable binaural audio recording, headphones	None	Traditional theater
Seance-flight-coma http://www.darkfield.org	Audience in a set in pitch darkness, props with tactile feedback responding to audio.	Pre-recorded binaural audio, complete darkness and props with haptic feedback.	None	Headphones	None	Immersive set, darkness
The missing voice https://www.artangel.org.uk/project/the-missing-voice-case-study-b/	Participants are walking in London, following a story through a headset.	Pre-recorded binaural soundscape and audio guide telling a story in relation to real locations in London.	Pause and play	Portable audio device, headphones	None	Streets
Symphony of a missing room http://www.lundahl-seitl.com	Blindfolded participants are instructed through headphones, hands and lights to explore and perform a gallery space as an imaginary space.	Site-specific binaural audio, guiding voices, light and human touch.	Real actors	Wireless headphones	None	Gallery
Etiquette http://www.rotozaza.co.uk/home.html	Audience participants perform the entire piece based on audio instructions in a real environment in front of an unknowing audience.	Audio-guide to enhance a real environment, instructing participant performers to perform a play.	One other participant	Headphones	None	Coffee shop
Binaural dinner date https://zu-uk.com	Participants are going on a real date and are guided via headphones to perform for each other.	Real-time binaural audio instructions to guide participant performers reacting to real-live events.	One other participant, observation and individualized content directed by observer	Wireless headphones	Response to participants actions by a real person	Restaurant
Ulrike and eamon compliant https://www.blasttheory.co.uk/projects/ulrike-and-eamon-compliant/	One person at a time with a mobile phone takes on the role of one character while walking through a city and interact with the story through phone calls.	Site specific audio delivered via mobile phone. Interaction via dialling numbers, recording answers, hanging up.	Phone calls, dialling numbers, recording the own voice	Mobile phone	None	Streets
Rider spoke https://www.blasttheory.co.uk/projects/rider-spoke/	Audience on a bike are directed with an audio guide and recorder through London. They are free to move but asked to record stories for other audience members to find and listen.	Audio guide for a cycle trip in London, participants stop and record their own stories on specific locations and listen to other people's contributions.	Location based audio, recording the own voice	Tablet, headphones	GPS	Streets
Cabin fever https://www.nytimes.com/2004/03/12/arts/art-in-review-janet-cardiff-and-george-bures-miller.html	Visitors wear headphones and look into a diorama of a village which is the backdrop for the binaural sonic narrative.	Binaural audio to enhance a miniature set that surrounds the visitor's head and can be explored by turning the head.	Turn the head	Headphones	None	Miniature set
Consequences https://www.playlines.net/shows	Visitors explore an immersive set with a location based narrative through bose frames. The audience moves freely and interacts with other performers.	Bose frames audio sunglasses, location based audio, interactive narrative, real live music performance.	Head movements, tapping the frames, interactions with other actors	Bose frames audio sunglasses	RFID tags	Immersive theater

2000; Zimmermann and Lorenz, 2008; Albrecht et al., 2011; Vazquez-Alvarez et al., 2012; Chatzidimitris et al., 2016; Hugill and Amelides, 2016), the potential of AAR for social interaction was evoked in early work (Lyons et al., 2000) and explored in studies focused on collaboration (Eckel et al., 2003; Ekman, 2007; Mariette and Katz, 2009; Moustakas et al., 2011). Many audio-only experiences are location-based, for instance in museums or

geocaching (Zimmermann and Lorenz, 2008; Peltola et al., 2009). Sonic interactions remain a challenge in gameplay, but recent studies looked at interactive audio in the games AudioLegend (Rovithis et al., 2019) and Astrosonic (Rovithis and Floros, 2018). Presence also remains under-explored in AAR, though it is widely examined in virtual reality (Slater and Wilbur, 1997). AAR allows asking “how immersive is enough” (Cummings et al., 2015) to

achieve a feeling of presence, since even with technology which is not fully immersive, participants can feel immersed and engaged (Shin, 2019).

Interactive digital technologies provide new ways for artists, actors and participants to enact their ideas and tell their stories (Galindo Esparza et al., 2019). But so far, few experiences make use of the interactive potential of AAR. **Table 1** gives an overview of the state-of-the-art of interactive audio technologies used in the context of theater and performance. Spatial audio and formats of AAR are used in immersive theater to create imaginary spaces and parallel realities¹ (Klich, 2017), and to engage audiences in different ways with the potential to prompt interactions, participation and collaboration². Headphone Theater, where the audience wear headphones, relies on a sonic virtual reality with binaural sounds which enforces immersion and enhances a multimodal experience of a performance (Todd, 2019). Many headphone theater performances are criticized to isolate the audience members from each other and not acknowledge the physical headphones as a material in the performance³ (Klich, 2017; McGinley, 2006). Both the National Theater performance ANNA⁴ and Complicite's The Encounter⁵ use binaural sound recording equipment live on stage to deliver a more intimate performance to the audience through headphones. Only few performances play with interactive audio technology (Gibbons, 2014; Wilken, 2014), Consequences by Playlines (2019)⁶ being one of the first immersive theater performances delivering site-specific interactive narrative combining both augmented sounds and real-life interactions.

This paper investigates the potential of AAR for multiplayer embodied interaction through participatory performance. By exploring the user experience in this context, it offers insights about AAR providing new kinds of actions, interactions, and ways to make meaning for audience participants. It addresses two research questions:

- How can audio human-computer interaction prompt and support actions in AAR experiences?
- How can AAR contribute to developing new kinds of interactions in participatory performances?

2 MATERIALS AND METHODS

This section presents the design and development of the participatory performance Please Confirm you are not a Robot, details the AAR architecture that was developed to

deliver it, and discusses the methods used to conduct the user experience study and answer the research questions.

2.1 User Experience Design and Storytelling

AAR experiences often assign a role to the user, asking them to perform or take on a character role in order to play the game. Audience participation is used as a medium to communicate different layers of a story. Participants become actors as well as the audience for each other. There is room for an audience to collaborate and communicate, bringing their ideas and experiences to the story. While theater often makes use of new technologies to create new experiences, AAR technology has not yet been used to its full potential. This section introduces participatory and immersive performances as material for development of the experience.

2.1.1 Immersive Theater, Participatory Performance and Their Technological Mediation

Immersive theater is a re-emerging theater style, based on the premise of producing experiences of deep involvement, which an audience commits to without distraction. Every performance which acknowledges the audience has the potential to be immersive. Immersive theater often hinges on architecture, found spaces or installed environments for audiences to explore or perform in, with the audience members' movement as part of the dramaturgy of the work. Immersive theater often involves technological mediation. In the context of theater and performance, Saltz (2001) defines interactive media as sounds and images stored on a computer, which the computer produces in response to an actor's live actions. The control over or generation of new media can be designed based on the live performance. Bluff and Johnston (2019) make an important case for interactive systems to be considered as material that shapes performance.

AAR as an immersive medium engages the audience in different ways (White, 2012). Audience participation in performances is often motivated to restructure the relationship between performer and audience. More active audience engagement gives the audience agency to co-create the story. Participatory performance techniques are at play in immersive theater as soon as the audience-performer roles are altered and the audience is somehow involved in the production of the play, going as far as overcoming the separation between actors and audience. Pioneers in participatory performance argue for the political potential of collective performance-making (White, 2012). Participatory performance allows the participants to explore identities, social situations or even imagine new ways of being (Boal, 2002).

The community-based theater practice Theater of the Oppressed (Boal, 2002) blurs boundaries between everyday activities and performance. This practice is used by diverse communities of actors and non-actors to explore and alter scenarios of everyday encounters, resolve conflict or rehearse for desired social change and co-create their community (Kochhar-Lindgren, 2002). Acting together can build cultural stories, foster community dialogue about unresolved social issues, exploring alternative responses. It is a practice of creating,

¹<https://www.nytimes.com/2004/03/12/arts/art-in-review-janet-cardiff-and-george-bures-miller.html>, <http://www.lundahl-seitl.com>

²<https://www.fanshen.org.uk/projects>, <http://www.internationalartsmanager.com/features/reality-theatre.html>, <http://www.rotozaza.co.uk/home.html>, <https://zu-uk.com>

³<http://www.darkfield.org>

⁴<https://www.nationaltheatre.org.uk/shows/anna>

⁵<http://www.complicite.org/productions/theencounter>

⁶<https://www.playlines.net/shows>

manipulating and linking spontaneously emerging acts in order to transform images and space, and subsequently identity and politics. The Theater of the Oppressed establishes a venue where alternative strategies, solutions and realities can be tried on, examined and rehearsed. The performance for this study was developed based on the games described in Boal (2002) since they are suitable for non-actors and provide tools to create a collaborative group experience.

2.1.2 “Please Confirm you are not a Robot”

The performance piece *Please Confirm you are not a Robot* is an interactive story that attempts to interfere with group dynamics. The narrative draws attention to the ways in which digital devices and services act negatively on the user’s mental and physical wellbeing. In a series of games, choreographies and exercises, participants are prompted to reenact, observe and subvert their digital practices (Ames, 2013; Galpin, 2017).

Please Confirm you are not a Robot revolves around a speculative fiction, constructed of four individual games. At the start, each of four participants meets their guide called Pi who introduces the participants to the scenario and subsequently guides the players through each sequence. Pi has a sonic signature in spatial audio and appears talking around the participants’ heads. Pi also introduces the interaction controls of the Bose Frames Audio Sunglasses: nodding, shaking the head and tapping the arm of the frames. Each game requires gradually more human interactions, always stimulated and guided by spatial audio and the narrator. The fictional goal of the training is for participants to deactivate cookies that have been installed in their brains. As the story evolves, Pi displays more and more human features and reveals their own feelings and fears individually to each participant, until Pi is trying to find their voice through one of the participants, prompting one player to speak out loud and end the performance by taking each player’s frames off.

2.1.2.1 *Circles and Crosses*

In the first game, participants are asked to draw a circle and a cross in the air with both arms at the same time. This is a warm-up exercise for people to get comfortable moving in front of each other and using the whole body to perform. Participants receive spatial audio cues of a chalk drawing a circle and a cross on a board in front of them.

2.1.2.2 *Mirroring*

For the second game participants pair up. Each pair is asked to mirror each other’s movements and gestures. While they are moving, they are prompted to ask each other questions and start a conversation. The study looks at whether this contributes to affect between participants, as well as how the different layers of sound are prioritised.

2.1.2.3 *Notification*

This game uses the Frames as a gaming device. It is the section with the most spatial sounds. Players hear various notification sounds around them in space. They have to turn toward the source. Once they face it correctly they hear a feedback sound and can switch the notification off by double tapping the arm of the

Frames. The first player to score 10 points wins. This game allows the study of interaction with spatial audio and audio feedback.

2.1.2.4 *Tapping for Likes*

Participants are prompted by Pi to give likes to the other players by tapping their Frames. At some point the likes appear in a random order, different for each participant, so participants don’t always know what they are liked for. The game requires participants to get closest to each other and enter into each other’s personal space. One player is asked to step away and watch the game while Pi opens up about their feelings. Pi asks the player to speak for them to the rest of the group. Subsequently the excluded player takes everyone else’s Frames off and the game is over. This last game requires one participant to take on a pre-described role and repeat a script. Successful completion shows that the players trust the voice and the game and allows the study of the performative potential of the technology.

2.2 Audio-Augmented Reality Architecture

This section presents the modular AAR architecture used to deliver interactive narrative for the user experience study. It starts by looking at the technological choices and then details the setup of the multiplayer gameplay.

2.2.1 Technological Choices

The Bose Frames Audio Sunglasses feature acoustic transparency, headtracking and user interaction possibilities (nodding, shaking head, and tapping the frames). Their acoustic transparency works by leaving the ears free to perceive the real auditory environment but adding spatial audio via little speakers in the arms of the frames. In order to make the AAR system modular, the project was developed in Unity. The prototype was developed for iOS phones only for compatibility with the Frames SDK. The Frames feature a low energy Bluetooth system. The headtracking uses an accelerometer, gyroscope and magnetometer, with latency of around 200 ms. Even if this is higher than the optimal latency of 60 ms, which may affect audio localization performance (Brungart et al., 2006), other aspects of the Frames, such as not covering the ears and being less bulky than other headphones with headtracking, made this device suitable to achieve good user engagement in the game. Google Resonance Audio SDK was used for 3D audio rendering because of its high-quality with 3rd-order ambisonics (Gorzel et al., 2019). **Figure 1** is a representation of this AAR architecture.

2.2.2 Multiplayer Gameplay

A Local Area Network over WiFi using Unity’s UNet system, allows a connected multiplayer experience with all players connected via a router. The first player who connects to the game is the host, meaning that the corresponding phone is a server and client at the same time. Subsequent players connect by entering the host IP address on their phones to join the game.

Each client owns a player whose movements correspond to the ones of the Frames. Some global elements are synchronized over the network such as the scores, while others are not synchronized and are used to trigger audio events for specific players. This

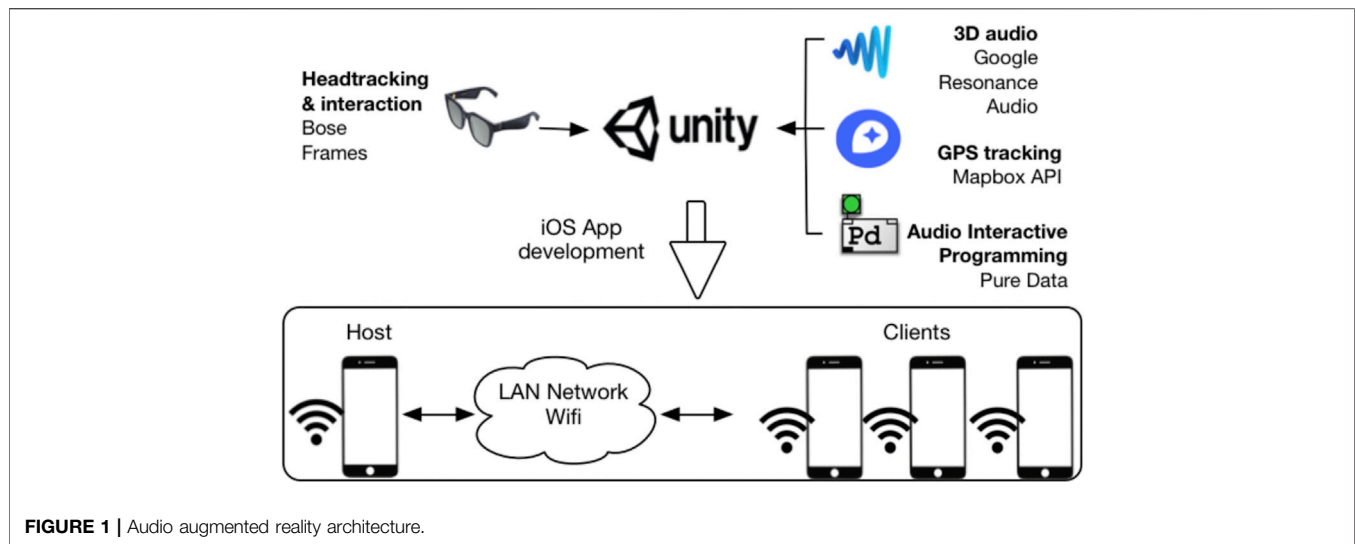


FIGURE 1 | Audio augmented reality architecture.

system allows different players to listen to different sounds synchronized over time.

The participants rely solely on the audio content and never have to look at the screen during the experience. For this prototype the only necessary user interactions with the screen consist in clicking on buttons to access each game from the main menu, start the multiplayer connection with the Start Host button, enter the host IP Address and then press Join Game, or quit the game. When each user finishes a game they automatically return to the main menu.

2.3 Study Design

This section details the research method used to carry out the user experience study. It outlines the test protocol and explains the analysis methods. This study was approved by the Queen Mary Ethics of Research Committee with the reference QMREC2161.

To answer research question 1) (“How can audio human-computer interaction prompt and support actions in AAR experiences?”) this study explores the perception of sound location considering the technology used, spatial sounds prompting movements, the participants’ engagement with spatial audio and their interest in it. The relative importance of different layers of sound is also considered.

For research question 2) (“How can AAR contribute to developing new kinds of interactions in participatory performances?”) the study looks at communication between participants and the effects of asymmetric information in AAR on a group. It furthermore investigates performative aspects of the experience as well as the potential of sound and movement for storytelling. A point of interest here is whether the experience can create a feeling of connection between people and thus create empathy for the protagonists of the performance.

2.3.1 Protocol

A pilot test with four participants from within BBC R&D was conducted in May 2019 to detect technical and narrative flaws and gather user feedback to refine the prototype. After

adjustments to the experiment design based on feedback from the pilot study, four groups of four participants each, with different levels of previous experience in AAR and performance, participated in the study in June 2019. Each session lasted 1.5 h. They were conducted on two consecutive days in the canteen of BBC Broadcasting House in order to test the experience within a naturally occurring soundscape.

Two of the researchers were in charge of manipulating the phones to start and change the games. This allowed the participants to only focus on the audio content and not pay attention to phones and screens. They also helped the participants if they had questions.

The evaluation was structured as follows:

1. Participants fill in a consent form and pre-study questionnaire
2. After each game, participants answer specific questions in a post-task questionnaire
3. At the end of the experience, participants fill in a post-study questionnaire and then take part in a 15 minute guided discussion

2.3.1.1 Before the Performance

The pre-study questionnaire contains 11 questions. Three questions concern demographics (age, gender, field of work). Seven questions address participant experience in spatial audio and AAR (3 questions), immersive theater and group performances (2 questions), use of a smartphone and social media (2 questions). Finally, one question concerns the feeling of connection of the participant with the group using the “Inclusion of Other in the Self (IOS) Scale” (Aron et al., 1992).

2.3.1.2 During the Performance

The post-task questionnaire asks between two and four questions at the end of each game about level of engagement and participant’s feelings. It aims at getting direct feedback that the participant could have forgotten after the experience.

Observations about the behaviors of the participants were carried out by the first two authors of this study, supported by one or two additional researchers who took notes during each experiment. Each observer was paying attention to one or two participants at a time. Each session was video recorded in order to compare the results.

2.3.1.3 After the Performance

The post-study questionnaire contains 21 questions about engagement, interactive ease, problems, storytelling and feelings about the group. The first six questions address user engagement, with overall feeling, and liked or disliked elements. Then, three questions focus on the digital authority, with two questions about the perceived authority of the voice narrator, and one question regarding the self-consciousness of the participant during the experiment. Following seven questions examine the feeling of presence, chosen from the IPQ presence questionnaire, which provides a good reliability within a reasonable timeframe. After that, four questions aim at getting insights regarding the game design. They address the user experience in terms of pace, interactions usability and clarity of the explanations. Finally, the last question examines the feeling of connection with the group after the experiment, again using the IOS scale (Aron et al., 1992).

The guided group discussion at the end of each performance contains eight questions about possible contexts and applications for technology and performance, audio-only experiences and 3D audio, asymmetric information, general thoughts about the story, reflective aspect of the experience around social media the overall feeling about the experience.

2.3.2 Analysis

The analysis was conducted using both qualitative and quantitative methods.

The qualitative analysis compared participants' answers from the questionnaires to their behaviors during the experiments, using both the recordings and notes taken by the researcher. Content from the guided group discussion together with the open ended questions from the questionnaires were analyzed using grounded theory (Glaser and Strauss, 1967). Answers were organized into meaningful concepts and categories, and progressively refined throughout the analysis process of the different questions.

Quantitative analysis was made using the pre-study, post-task and post-study questionnaires, counting the number of answers which appeared. Comparison of pre- and post-study questionnaires generated insights. In particular the feeling of individual inclusion within the group (using the IOS scale) and the performance of participants in relation to their previous experience in the fields of 3D audio and immersive performance.

3 RESULTS

This section first presents the demographics of the participants and then describes the results and analysis of the investigation regarding the questionnaires, semi-directed interviews and behavioral observations.

3.1 Demographics

Participants were recruited from the researchers' personal and professional networks (BBC, Queen Mary University London), using social media (e.g., Facebook, LinkedIn) and mailing lists. In order to get a more even number of female and male participants, a mailing list for women in technology called "Ada's List" and the Facebook group "Women in Innovation" was targeted. Nevertheless, the recruitment resulted in a lack of diversity within the groups. Due to the large time commitment required from the participants (1.5 h) it did not seem appropriate to be more selective with recruitment.

Participants in the pilot study were all BBC R&D staff, aged from 22 to 40, 1 female, 3 male. Three of them had previous experience in listening to 3D audio and binaural audio, but nobody in AAR. One participant had a previous experience in immersive theater and one in performance. The actual study contained 16 participants, aged from 24 to 40. 11 male and 5 female participants. All participants worked in one of the following fields: engineering, creative industries, research, technology, audio or entrepreneurship. They all used smartphones and social media in various different contexts. 14/16 participants had listened to spatial audio before participating in the study, i.e. binaural audio for 12/16 participants, and 10/16 had experienced at least once an AAR or audio only experience. Half (8/16) of the participants had visited immersive theater plays or performed in front of an audience themselves.

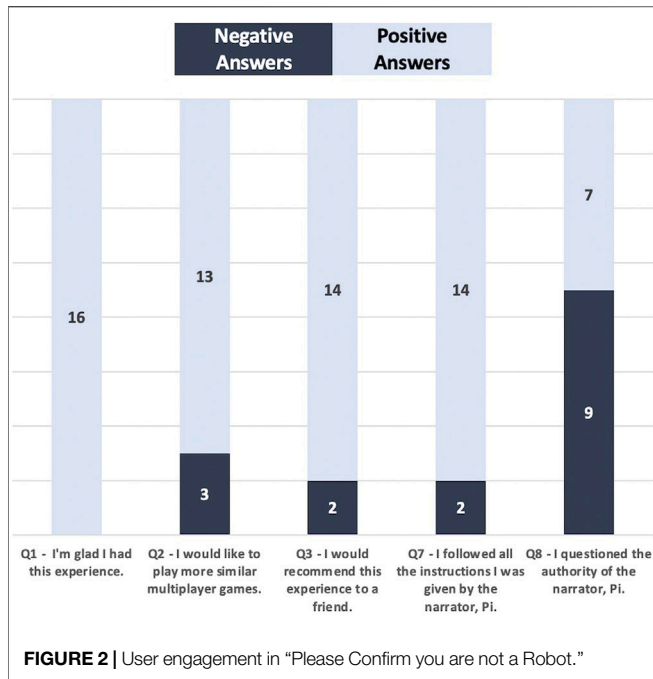
3.2 Questionnaires

Analysis of the questionnaire results was carried out using quantitative methods and grounded theory. Where possible, comparisons are made with the profiles of the participants. The questions can be grouped into two key areas of interest: user engagement (including multiplayer AAR experience, cognitive load and AAR design) and the feeling of presence. The results are presented below along these categories since they emerged from analysis across all methods used for the evaluation of the experience.

3.2.1 User Engagement

The questions 1–8 of the post-study questionnaire (see **Supplementary Material**) target the user engagement with the experience and the performance. The results to the questions 1–3, 7 and 8 are presented in **Figure 2**. Participants had four possible answer choices: Don't agree at all, Somewhat disagree, Somewhat agree, Totally agree, followed by an open question to explain why they chose the answer. For clarity the two positive answer categories and the two negative ones have been summarized in **Figure 2**. The answers will be discussed in the following subsections.

Figure 3 summarizes the categories and concepts that emerged from the grounded theory analysis of the open ended questions. It presents what the participants like and dislike about the experience and performance. The number of phrasings for each related category is represented by a number inside parenthesis. Well received are the social and embodied aspects of the performance, forming the biggest category (18). Even if the



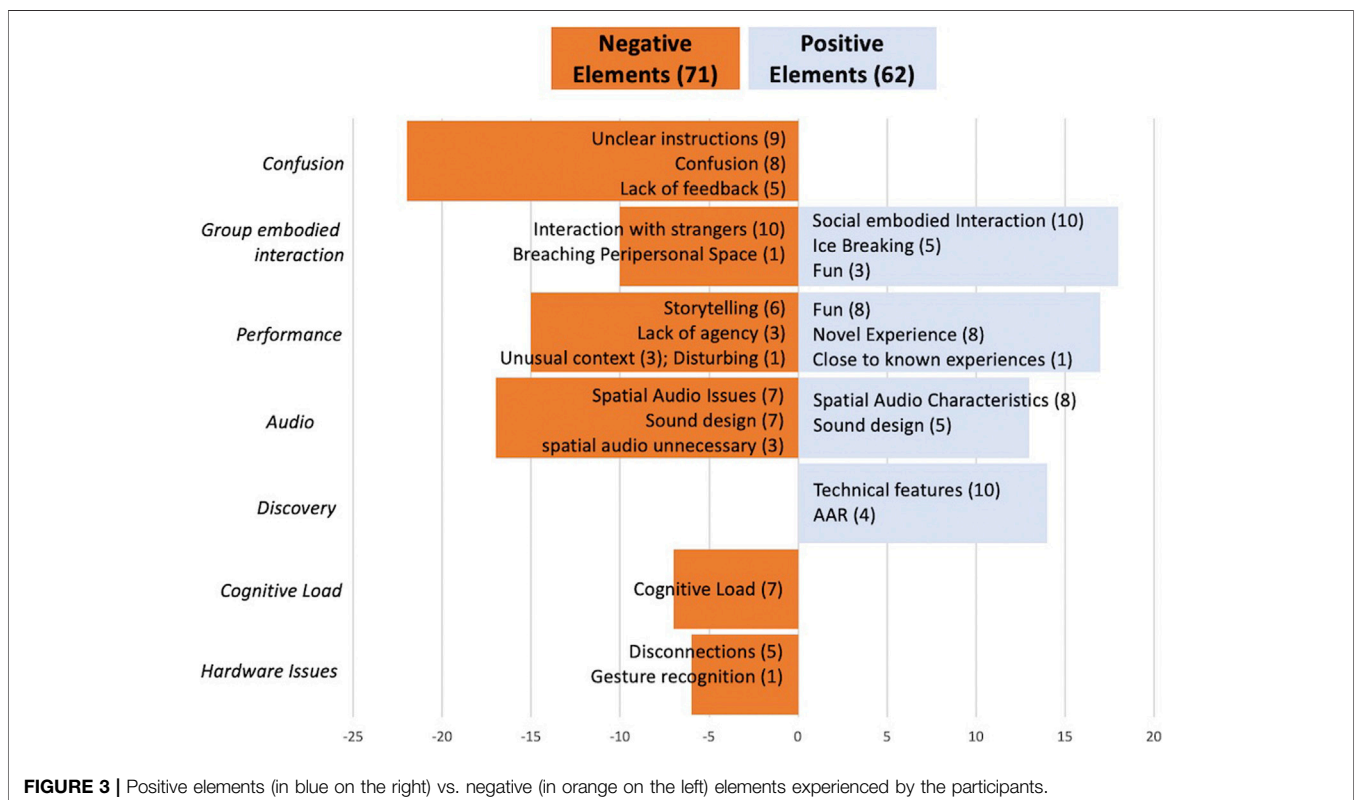
interaction with strangers is at times perceived as uncomfortable (10), the experience invited participants to interact with the group using their whole body (10), “relying on my body language” (P13), and worked as an ice-breaker (5). The experience was fun

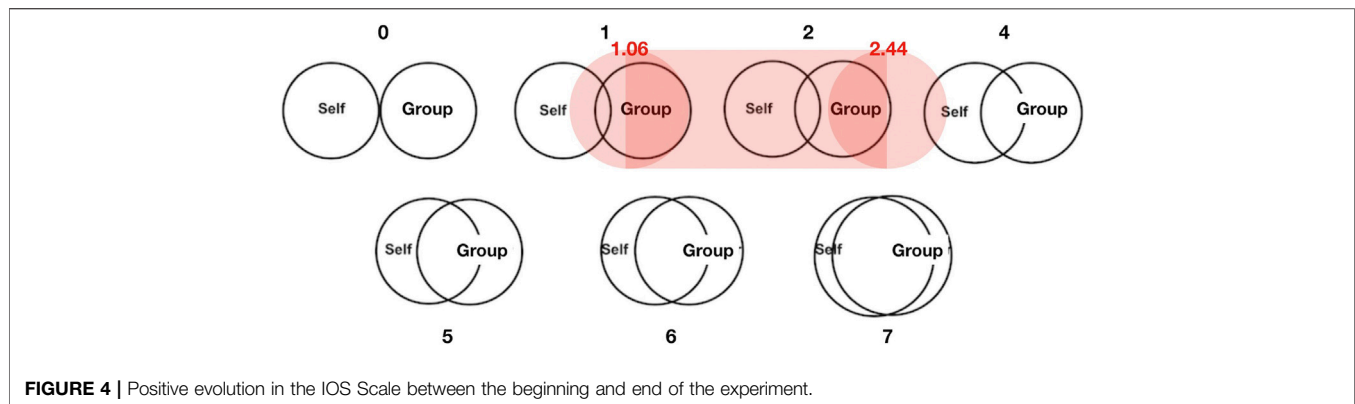
(8) and novel (8), and the discovery of new features of the AAR technology (14) contributes to its enjoyment. Though, participants often criticize the storytelling (6) or their lack of agency (3) within the performance. Audio is a mixed category, with almost as many positive as negative statements, regarding the sound design as well as the spatial audio features. Many participants were confused, mainly due to unclear instructions (9) and lack of audio feedback (5), often leading to cognitive load (7).

3.2.1.1 Multiplayer AAR experience

As shown in **Figure 2** all participants enjoyed taking part in the performance. A key reason for enjoyment is an interest in the social interaction as a group. Most participants (13/16) would like to play more similar multiplayer games or performances in the future. The novel interactions within the performances made possible by the AAR technology invite participants to “push social boundaries” (P13). It is deemed very suitable for ice breaker activities and provided a platform for strangers to interact with each other.

The results of the IOS scale in question 22 clearly show that the feeling of closeness to the group, even if temporarily, increased during the experience. Every participant showed an increase of at least one point (see **Figure 4**). Across all groups, participants started the study at an average of 1.06 on the scale and ended the experiment at an average of 2.44. This result was backed up by observations that the participants made during the discussion who started shy, hesitant or intimidated and then opened up and became more relaxed (P5, P7): “I liked that it brought a group of





complete strangers to interact together.” (P7); “Performing for others I didn’t know, but that was kinda nice.” (P8).

3.2.1.2 Cognitive Load

Participants do not always follow the instructions they receive. The reason for this is mostly some sort of confusion, misunderstandings of the story or overwhelming cognitive load. This happens when the audio content is delivered while participants are already carrying out a task. Particularly in game 2, “Mirroring,” 8/16 participants self-report facing too many concurrent tasks, or being distracted or interrupted by the audio while trying to finish a task. 5/16 participants ignore several of the requests to speak, or ask questions. “Instructions were sometimes long and you can’t get them repeated if you zone out.” (P9, P10, P11, P12), “I did not like when a talking prompt interrupts a conversation.” (P10, Group 5).

Participants do accept the instructions as the rules of the game and try to follow them as much as possible. Most participants (9/16) trust in the narrator and accept instructions as game rules, considering that they have no choice than to follow. The authority of the narrator is questioned where the instructions are unclear.

3.2.1.3 AAR design

Audio sequences with strong spatial audio cues are well received and contribute to the enjoyment of the experience. The audience is favourable toward the spatial audio where it appears close to the real environment. This contributes to a feeling of immersion and was facilitated by the acoustic transparency. Spatial audio experts did not comment on the audio.

During the game “Mirroring” several sounds which could be mistaken for real-world sounds were introduced (snigger, door slam shut, mosquito) to observe whether or not participants react to them. Only one participant (P16) reacted by imitating the mosquito.

Regarding spatial sound localization in the “Notification” game, 11/16 participants say it was unclear to hear where the sound was coming from in terms of azimuth and elevation, as shown in question 5 of the post-study questionnaire, and head movements did not help. 8/16 participants express frustration:

“There was some relief but frustration was more not being able to locate them. Sharp left right was obvious but nothing else was.” (P6).

3.2.2 Feeling of Presence

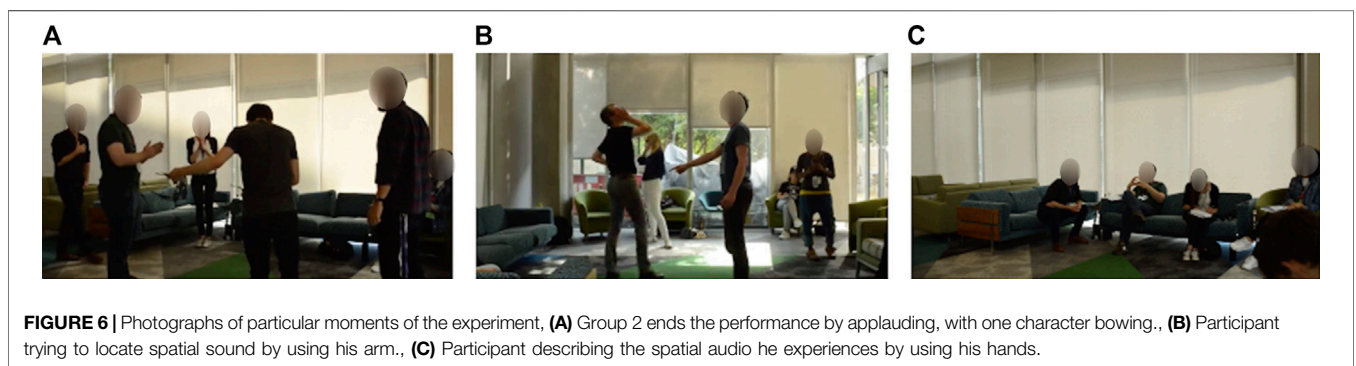
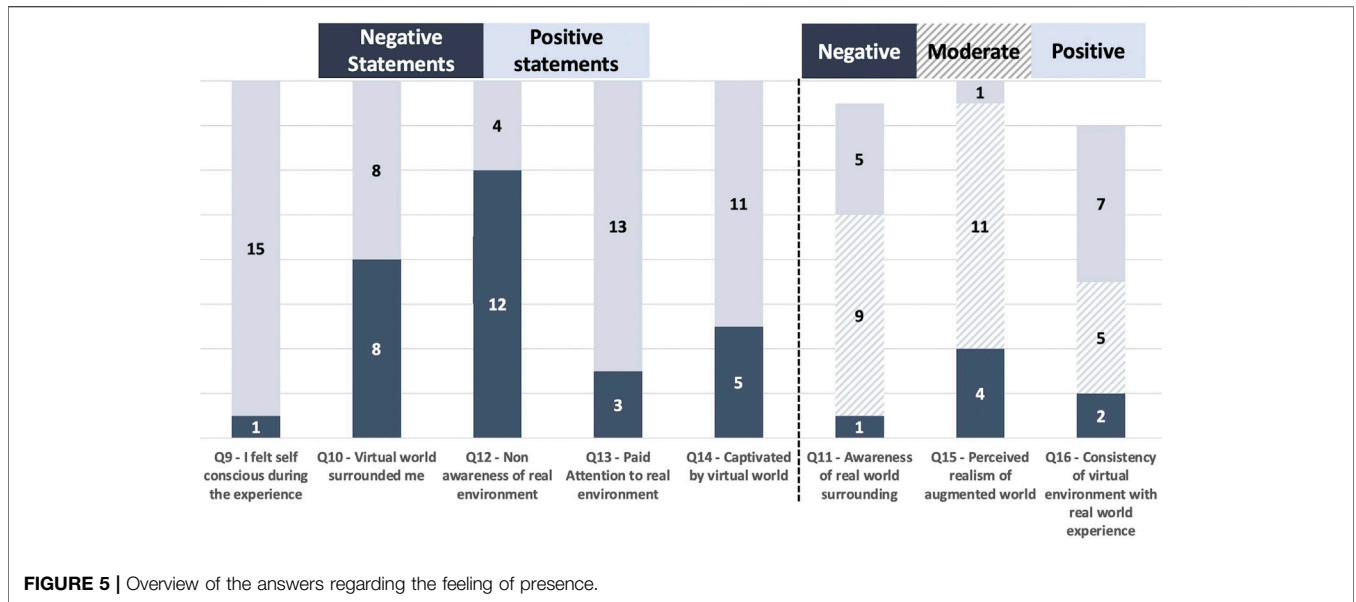
This section summarizes the results of eight questions addressing the feeling of presence. 6 of them were taken from the IPQ presence questionnaire (Schubert, 2003). **Figure 5** presents an overview of the questions and numerical results, with on the left the four-answer scale questions (Q9, Q10, Q12, Q13, Q14), displayed as positive and negative statements for the sake of clarity, and on the right the three-answer scale questions (Q11, Q15, Q16).

Nearly all participants (15/16) felt self-conscious during the experience because of the required interaction with strangers (6), uncertainty (3) and the unusual situation (2). The interaction with strangers concept is echoed in the semi-structured interviews, where participants said for example: “It was weird to interact with complete strangers. There was no basic small talk to break the barriers.” (P15, P17); “Maybe it would be more enjoyable to play with friends.” (P18, P20).

Half of the participants (8/16) somehow felt that the virtual world surrounded them. 11/16 participants were completely captivated by the augmented world and 12/16 perceive it as moderately real. The fact that there is only audio and the participants constantly see the real world makes the augmented world less real (6). Where participants don’t perceive the augmented world as real is partly due to the sound design (2).

Almost all (14/16) participants remained aware of the real world surrounding them while navigating the augmented world. This due to the fact that they are watched (P17) and “aware of people seeing me doing weird things” (P15), and also because there are no visuals other than the real world. Participants thus pay attention to the real environment (13/16).

7/14 participants did not find the virtual experience consistent with their real world experience. This is mostly attributed to audio issues (10) and the particularity of the experience (4).



3.3 Semi-Directed Interviews and Behavioral Observations

After the experiment each group was asked to discuss the AAR experience in guided interviews. Answers are grouped into the topics that emerged from this discussion and complemented by observations made during the performance.

3.3.1 Interactions

Participants liked the game interactivity, and the most liked game (i.e., Tapping for Likes) was indeed the most interactive one. 14/16 of participants felt comfortable or very comfortable with other participants touching their frames. The Frames were considered an object that is not part of the personal space, even though they are very close to the face.

Participants started to perform and take on a role during the experience. 5/8 pairs verbally coordinated their actions and discuss what scenario to improvise when prompted to move synchronously in the Mirroring game. At the very end, participants tend to become performers. In particular, the excluded participant of most groups follows but alters the script. “And now I have to take you all back to

reality, by taking your glasses, and then you’ll truly be free. So that’s how it ends.” (P12) P12 ends the experience by giving a bow. Group 2 applauds when the experience is over (see **Figure 6A**) “It’s been a pleasure!” (P12).

Participants perceive the non-verbal interactions mediated by the technology (i.e., tapping the frames or mirroring) as strange compared to the real world. They copy each other’s behaviors to understand how to play the game. The interaction with the audio content was linear and static, which is disruptive to the performance of participants since they have no control over repeating the content - a concern mirrored in the answers to Q7 and Q8 of the post-study questionnaire.

The audio content was personalized but did not clearly reflect actions or interactions of other participants. Participants would have preferred to know if they were synchronized in their information stage. “If I got tapped I could hear it. If I tapped somebody I had no feedback it had been successful.” (Group 5) Few participants were confused and uncomfortable when the instructions were not clear.

3.3.2 Audio

Since the experience included asymmetric information, all groups discussed that they would have preferred more information about the discrepancies in information they receive. Participants suggested to include sonic signature to differentiate between private and shared content or a disclaimer when the content is about to differ. People worked together to figure out that they hear different information: “Oh, for me it’s different.” (P10, P12), “Did you hear that? It told me to double tap when we’re ready. Are you?” (P9, P12). Some participants tried to signal with gestures to the rest of the group that they received different instructions. Few participants were positively surprised when they found out that they heard different things. The potential of asymmetric information through individual devices can be further explored. “When at some point I realized we got different instructions I thought it is really great.” (Group 3).

The participants do not want to miss out on information from the narrator. When they are listening to the narration they all usually stay very still and focused. They try to stop other participants from interfering while they are listening and they do not like talking over the voice while listening. They use non-verbal communication instead, if they have to communicate with the group. P8 gives a thumbs up to show that they are connected rather than replying by answering.

The audio only aspect of the experience was received positively, allowing participants to focus on group collaboration and on each other. Adding more technologies or visuals was by some perceived as too much: “I don’t think it needs another technology on top, the world is enough of a screen.” (Group 3). Sounds make the participants visualize things in their head, “for one it is a dot, for another it is like an actual phone.” (Group 4) Though, fewer participants suggested adding physical props to the experience, such as inanimate objects or other technologies. Only two participants would have preferred to have visuals to feel more immersed in the experience.

Participants report two types of spatial audio issues, i.e., sound positioning and integration between real and virtual sounds. They had trouble locating the audio sources in the Notification game, including spatial audio expert users (P5, P6, P8, P10, P18), and noticed head-tracking latency, reporting that sounds “jumped somewhere else” or “were moving”. Despite integration issues, the spatial audio was particularly impressive when it was not clear whether sounds were augmented or real: “I could hear someone who dropped a fork but could not tell if it was coming from the glasses. I thought it was pretty cool.” (Group 5). Participants would have preferred more spatial audio elements throughout the experience.

Observations and group discussion show that people look for and identify spatial audio cues horizontally on eye-height as well as upwards, but rarely below them. One participant closed his eyes and tried locating the sound by using his arms to point at the direction he heard it coming from (see **Figure 6B**). Observation during discussion showed that hands are actively used to describe spatial sounds (see **Figure 6C**). Participants

who had never listened to spatial audio before expressed an interest in a dedicated onboarding at the start of the experience. During the notification game 9/16 participants moved around in space to locate the sounds, using their ears to navigate, and the other 7/16 participants remained static, turning their head around. “Fun to use ears to navigate.” (P8).

3.3.3 Other Design Considerations

Participants liked that their ears were not blocked and thus able to hear their surroundings. Yet, the format of glasses was perceived as strange since the lenses are not used and glasses are usually worn for specific purposes not relevant to the given experience. The darkened glass made it impossible for participants to look into each other’s eyes. Speakers could be inserted in other head-worn objects such as a hat.

Lack of clear feedback sounds was discussed by all groups. The existing feedback was confusing as it remained unexplained. The feedback sounds were not clearly distinguishable from sounds which are part of the content or game. The hardware itself does not give any tactile feedback. “My actions didn’t get any feedback unless I was correct and often then I felt like I was guessing.” (P6, P8). There was no feedback response on how strong to tap so that it works but without hurting the other.

The most popular idea for other interactive AAR applications was experiences tied to specific locations via GPS, exploring a city or space (e.g., a museum) in form of a story or guided tour. The interactive technology would be popular for party games or to enhance board games. The participatory and collaborative aspect of the experience was deemed especially suitable as ice-breaker activities, warm ups, workshops or speed dating. Other applications could be seen in acting classes or to support the visually impaired.

3.4 Evaluation of “Please Confirm you are not a Robot”

All but two participants (14/16) say they would recommend the experience to a friend since they enjoy the novelty of the technology (4) and the story (3). Technological issues with hardware and audio (3) as well as confusion about instructions and the story (9) had adverse effects on the enjoyment for the participants. The storytelling was sometimes described as simplistic (3) and boring (2) and spatial audio, despite being liked, not necessary for the story (2).

Two games were preferred, i.e. Tapping for likes (8/16), and Mirroring (7/16). In the Tapping for likes game 10 participants found it exciting to get likes. Yet, the aspect of asymmetric information for giving likes seemed arbitrary and took away from the participant’s perceived agency. They preferred when there was consensus. The Circles and Crosses and Notification game each got one vote. Circles and Crosses was an individual exercise but with the other participants as spectators. It worked as an icebreaker since 12/16 of participants found it difficult to achieve but more than half 9/16 still laughed and had a fun time.

TABLE 2 | Main Findings.

RQ1: Audio HCI	RQ2: AAR in participatory performance	Other findings
<ul style="list-style-type: none"> • Asymmetric information potential for group coordination • Simple gesture interaction potential for coordinating the storyline • Allow for more control over the sonic content • AAR can make users feel “present” 	<ul style="list-style-type: none"> • Real-time audience data can be used to generate the narrative • Spatial audio directs and impacts performers’ movements • Interactions with audio content give the performers control over the story • AAR allows for embodied experience and empathy with the group 	<ul style="list-style-type: none"> • Ice-breaking effect and long last impact • Creative freedom for group personality to form • On-boarding needed for spatial and feedback sounds • Participants rarely look for spatial audio cues downwards

4 DISCUSSION

The objective of this experiment is to examine how audio interactions can prompt actions in AAR experiences and how AAR supports interactions in participatory performances. This section discusses the results from the user experience study to answer the research questions. Other findings are also discussed that introduce areas for further research. **Table 2** summarizes the main findings of this study.

4.1 How can Audio Human-Computer Interaction Prompt and Support Actions in AAR Experiences?

Asymmetric information, by allowing different users to hear different parts of the story, prompted group coordination through verbal and non-verbal communication. Participants who discovered that they were hearing different elements gave accounts to each other of what they heard, discussed how to solve tasks or perform as a group. Thus, asymmetric information had a strong multiplayer potential and prompted discovery. Yet, even if some participants really enjoyed it, others got confused and criticized the lack of information about the discrepancies in information they received. The analysis of the post-study questionnaire and discussions reveals that confusion often comes from cognitive load, a common side-effect also noted in previous audio-only games (e.g., Papa Sangre.⁷). To prevent it, some AAR studies suggest finding the right audio-only interaction techniques (Sawhney and Schmandt, 2000), using minimal information, or more guidance as interactions become more complex (Rovithis et al., 2019). Hence, using an adaptive level of feedback in AAR depending on the user’s perceived challenge is advised. Otherwise, some participants express that using too little audio feedback is confusing, too challenging and increases cognitive load. To that respect, the meta-architecture exposed in the LISTEN project which provides meaningful audio descriptors helping to make context-aware choices regarding which sound to hear when during the game could be used (Zimmermann and Lorenz, 2008). Audio descriptors depend on the game-purpose: technical descriptions, relations between real and virtual objects, content of real objects, intended user etc.

Please Confirm you are not a Robot shows that in a multiplayer AAR experience asymmetric information could be used as an audio description to distinguish shared content from private information. This can be achieved by using different sonic signatures, or as in the performance *Consequences* (Playlines, 2019) by placing sound sources in the real world (for example actors, speakers or a live performance).

Some interaction design insights come from the analysis process. Double tapping the frames to start and finish every game was successful to coordinate the storyline as well as the participants as a group. Yet, participants often report a lack of control over the audio content. This may be due to hardware issues (non-detection, disconnection), the narrative moving too fast and interrupting the players when they were acting, or the impossibility to repeat content if it was not understood for the first time, nor to pause the narrator voice. Hence, offering repetition possibilities, and an audio system designed to not interrupt the participants could be further explored. Also, feedback sounds were sometimes not clearly distinguishable from sounds that are part of the story. As reported by most participants in the Notification game and in line with previous spatial audio studies (Brungart et al., 2006), the high audio latency of the Frames confused participants. They also had troubles finding the audio cues when located downwards, because they mostly looked for them horizontally as well as upwards, which supports previous results by Rovithis et al. (2019). They also mention that continuous spatial sounds (a continuous tone or music) are easier to locate than interrupted sounds (beeping). Hence, to facilitate interaction with AAR content a distinct on-boarding to audio controls, feedback sounds and how to listen to spatial audio is advised. Despite these challenges, participants were excited about the sound during the introduction and the parts where the spatial audio clues were strong, and suggest that they could play a bigger part in the story, or be integrated with a strong reference to the real environment, which is further explored in AAR studies part of the LISTEN project (Zimmermann and Lorenz, 2008).

Contrary to the existing AAR literature (Mariette, 2013), this study explores audio interactions through the angle of user engagement. Most participants were glad of the experience and liked its novelty, as in previous multiplayer AAR studies Moustakas et al. (2011), either because they tried a new technology or found the experience thought provoking. Most of them would recommend it to a friend and were engaged. One

⁷<http://www.nickryanmusic.com/blog/papa-sangre-3d-audio-iphone-game>

participant sent an email a few days after the experiment, describing how it made them reflect on their own technology use, revealing a longer term impact. The engagement of some participants led them to feel immersed in the environment. Even if they were still aware of their real surroundings, paid attention to it, felt self-conscious, they were captivated by the experiment and half of them felt that the virtual world surrounded them. Even if the augmented world was perceived as moderately real, it was consistent with a real world experience. These results show that participants experienced a form of presence in the story, for instance illustrated during the discussion where some participants found it impressive when it was unclear whether sounds were augmented or real. As suggested by Cummings et al. (2015), this result challenges the common assumption that a high immersion is needed for a good engagement and feeling of presence. It also supports previous evidence found by Moustakas et al. (2011), where high levels of perceived user immersion were observed in the Eidola AAR multiplayer game.

4.2 How can AAR Contribute to Developing New Kinds of Interactions in Participatory Performances?

Real-time data was used to coordinate storylines based on the choices and actions of all participants, thus creating a multiplayer experience. This has previously not been possible to achieve in real time, unless humans are present to observe the actions and direct the narrative as in the Binaural Dinner Date experience⁸, or the narrative evolves linearly, where one element always follows another⁹ or actions and recordings are tied to specific GPS locations and have to be placed there before an audience arrives¹⁰. In this experiment this group co-ordinated audio was explored in three out of four games. In *Mirroring* this allowed people to choose and pair up with another participant and receive the complimentary audio tracks in order to perform as a pair. In the *Notification* game each participant received an individual set of sounds to turn off, but the points they achieved are coordinated and a competition with a winner can be identified. Tapping for likes created a curious friction when the content first started in parallel but then diverged, making the interactions seem more absurd and eventually separated one participant completely from the group to perform a different role. This study shows that narratives can be emergent from the actions of the audience and recombined in ever different ways and forms.

Head-tracked spatial audio soundscapes, as opposed to static binaural audio, invite for exploration of the sound sphere. The experience used the sphere around the head of each user as the frame of reference for sounds to appear. Participants use their whole body to move into the direction of sounds and look for sonic cues. When sound sources are placed in a specific spot in the sonic sphere listeners use their arms and hands to locate or follow

the sound. Spatial audio, narrative pace and sonic cues impact the way the performers move. Further research on the embodied experience of audio augmented reality and spatial audio would be interesting to investigate relationship between sound and body movement.

During the study the researchers noted that each group performed and acted distinctively different as a group from each other. The AAR system guided the performance and the overall experience was perceived similar, nevertheless the freedom of movement and interpretation allowed by the system made it possible for each group to choose their own pace and own set of interactions. The mood of each group was coherent within the four games but separate from other groups. Each participant had the freedom to add or subtract from the content as they please. Once trust in the experience is built, the participants dare to engage more and create their own character which they express through gestures and enactment. This is in line with the discussion of White (2012) that AAR as an immersive medium is suitable for audience engagement in a performance. By allowing performers to interact with the narrative, the actors can set their own pace in which they want to move through the story. The more control over the audio each participant has, the easier it is for them to follow instructions. As suggested by Bluff and Johnston (2019) the AAR technology shapes the experience and the performance.

The best parts of the experience were described as the social interactions. Participants saw a potential in the games to break down social norms and barriers. It was enjoyable to physically interact with other people and most of them would like to play more multiplayer games. This insight supports previous results from Moustakas et al. (2011)'s AAR multiplayer game, where 80% of players enjoyed playing with a remote player. Thus, future AAR applications could focus on this multiplayer aspect, with participants evoking party games or board games. The results from the IOS scale show that collaborative action with multiple participants can create a feeling of connection. This positive feeling within the group could create empathy for the characters that emerge from the group performance since the performance received positive feedback for creating understanding around digital wellbeing through embodied, immersive engagement with a topic and narrative.

4.3 Limitations and Future Perspectives

The design of the experiment, questions asked about each task immediately after its completion, interfered with the flow of the performance and may have impacted on how participants perceived the experience. The testing situation presented an unnatural environment. Due to observers and by-passers being present at any time it was not easy for participants to let go. Furthermore, there was no way for researchers to monitor what participants heard at any given moment. This may have led to overlooking certain reactions and made it difficult to identify technological breakdown over human failure.

Hardware issues (e.g., connection failures, gestures not recognized) led to confusion and frustration and affected the overall perception of the experience. The researchers were not meant to interfere with the experience but due to technical issues observers had to intervene nevertheless. The high audio latency limited the spatial audio quality.

⁸<https://zu-uk.com>

⁹<https://www.emergencychorus.com/you-have-reached-emergency-chorus>

¹⁰<https://www.blasttheory.co.uk/projects/riderspoke/>

The Frames were criticized for being too bulky and not meaningful for the context of the performance. There was no reference in the experience to the participants wearing glasses and thus no acknowledgment of the technology, which made it seem arbitrary.

For future research and evaluation of the same experience it could be tested in one go rather than split into chapters. The order of the four chapters was defined by the narrative and could not be altered during the experiment. Switching the order of the games for each group might avoid possible bias. A follow-up study should reach out to a more diverse audience and ensure a better gender balance. The experience should also be tested in a more natural environment. Overall participants would have preferred more story-line as part of the experience. They did not feel as the character in a story but just followed instructions. There could be more narrative drive to hold it all together.

5 CONCLUSION

The Audio Augmented Reality (AAR) collaborative experience Please Confirm you are not a Robot was designed, implemented, and tested with the aim to evaluate the potential of AAR as a material to shape the content of a participatory performance through human-computer and human-human interactions. The experience was designed to explore the affordances of state-of-the-art AAR technology, namely asymmetric audio information between participants, haptic and gesture audio-control. The experience was tested and evaluated using questionnaires, observation and a guided group discussion. The results show that combining 3D audio with simple gesture interactions engaged the participants and made them feel present. Participants enjoyed the audio only aspect. The social aspect of the experience was the most liked feature. The audience, most often strangers to each other, communicated using verbal and non verbal interactions to perform tasks and navigate asymmetric information. The connection between members of each group increased over time.

Many previous AAR studies focus on interactive possibilities with devices and media with little exploration of user experience. Please Confirm you are not a Robot draws from performance practices to generate an embodied experience for the participants benefiting from real-time audio-interactions to generate narrative content. Interactive AAR utilizes real-time data for generating content and interacting with the media. Interactive, spatial audio expands previous experiments and performances with binaural audio for more personalized and immersive experiences. The option to generate narratives based on the status of other participants enables performances for multiple participants, each receiving a personal but collaborative experience. AAR has creative potential for the participant performers to express themselves and their mental images based on sonic stimuli. Interaction between participants—away from purely scripted plays—increases the feeling of empathy for the characters of a story and generates a higher feeling of being immersed in the story.

The use of AAR in participatory or immersive performance has not yet been explored to its full potential, but the results of this study show that the medium allow for novel ways of telling

stories. In future studies the use of real-time data could be explored in more detail of how it can generate and influence the story.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found at <https://doi.org/10.5281/zenodo.4457424>

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Queen Mary Ethics of Research Committee, Reference QMREC2161—Advanced Project Placement Audio Augmented Reality Experience. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AN worked on interaction and experience design, and story and script development. VB developed the audio augmented reality architecture, iOS App and multiplayer gameplay LAN over WiFi. Both AN and VB were equally involved in the study design, execution and analysis. For this paper VB focused more on Research Question 1 and AN focused more on Research Question 2. The paper was written in equal parts. PH, JR, HC, TC, CB, and CP were supervisors and consultants to the research project.

FUNDING

Funding support came from the EPSRC and AHRC Center for Doctoral Training in Media and Arts Technology at Queen Mary University of London (EP/L01632X/1) and BBC R&D.

ACKNOWLEDGMENTS

Thanks to BBC R&D IRFS and Audio teams, in particular Joanna Rusznica and Holly Challenger for helping with UX research; Joseph Turner for documenting the study and Thomas Kelly for technical support. Many thanks to all participants, without whom this research would not have been possible. Special thanks to Brian Eno, Antonia Folguera and Emergency Chorus for feedback on our project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frvir.2020.610320/full#supplementary-material>.

REFERENCES

- Albrecht, R., Lokki, T., and Savioja, L. (2011). "A mobile augmented reality audio system with binaural microphones," In Proceedings of interacting with sound workshop on exploring context-aware, local and social audio applications—IwS' 11, Stockholm, Sweden, August 15, 2011 (Stockholm, Sweden: ACM Press), 7–11. doi:10.1145/2019335.2019337
- Ames, M. G. (2013). "Managing mobile multitasking: the culture of iPhones on Stanford campus," In CSCW, San Antonio, TX, February 23–27, 2013
- Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Pers. Soc. Psychol.* 63, 596–612. doi:10.1037/0022-3514.63.4.596
- Bluff, A., and Johnston, A. (2019). Devising interactive theatre: trajectories of production with complex bespoke technologies. In Proceedings of the 2019 on designing interactive systems conference, San Diego, CA, June 23–28, 2019, (San Diego, CA: ACM), 279–289. doi:10.1145/332276.3322313
- Boal, A. (2002). *Games for actors and non-actors*. Hove, East Sussex, United Kingdom: Psychology Press, 301
- Brungart, D., Kordik, A. J., and Simpson, B. D. (2006). Effects of headtracker latency in virtual audio displays. *J. Audio Eng. Soc.* 54, 32–44
- Chatzidimitris, T., Gavalas, D., and Michael, D. (2016). "Soundpacman: audio augmented reality in location-based games," In 2016 18th mediterranean electrotechnical conference (MELECON), Lemesos, Cyprus, April 18–20, 2016 (New York, NY: IEEE), 1–6. doi:10.1109/MELCON.2016.7495414
- Cummings, J. J., Bailenson, J. N., and Fidler, M. J. (2015). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychol.* 19 (2), 272–309. doi:10.1080/15213269.2015.1015740
- Drewes, T. M., Mynatt, E. D., and Gandy, M. (2000). Sleuth: an audio experience. Atlanta: Georgia Institute of Technology, Technical Report
- Eckel, G., Diehl, R., Delerue, O., Goiser, A., González-Arroyo, R., Goßmann, J., et al. (2003). Listen—augmenting everyday environments through interactive sound-scapes. United Kingdom: Emerald Group Publishing, Technical Report
- Ekman, I. (2007). "Sound-based gaming for sighted audiences—experiences from a mobile multiplayer location aware game," In Proceedings of the 2nd audio mostly conference, Ilmenau, Germany, September 2007, 148–153
- Engel, I., and Picinali, L. (2017). "Long-term user adaptation to an audio augmented reality system," In 24th international congress on Sound and vibration, London, July 23–27, 2017
- Galindo Esparza, R. P., Healey, P. G. T., Weaver, L., and Delbridge, M. (2019). Embodied imagination: an approach to stroke recovery combining participatory performance and interactive technology. In Proceedings of the 2019 CHI conference on human factors in computing systems—CHI '19, Glasgow, Scotland, May 2019 (Glasgow, Scotland, United Kingdom: ACM Press), 1–12. doi:10.1145/3290605.3300735
- Galpin, A. (2017). Healthy design: understanding the positive and negative impact of digital media on well-being. United Kingdom: University of Salford, Technical Report
- Geronazzo, M., Rosenkvist, A., Eriksen, D. S., Markmann-Hansen, C. K., Köhler, J., Valimaa, M., et al. (2019). Creating an audio story with interactive binaural rendering in virtual reality. *Wireless Commun. Mobile Comput.* 2019, 1–14. doi:10.1155/2019/1463204
- Gibbons, A. (2014). "Fictionality and ontology," In *The Cambridge handbook of stylistics*, Editors P. Stockwell and S. Whiteley (Cambridge: Cambridge University Press), 410–425
- Glaser, B. G., and Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*. Paperback Printing Edn (New Brunswick: Aldine Publishing), 4
- Gorzel, M., Allen, A., Kelly, I., Kammerl, J., Gungormusler, A., Yeh, H., et al. (2019). "Efficient encoding and decoding of binaural sound with resonance audio," In Audio engineering society international conference on immersive and interactive audio, York, UK, March 27–29, 2019, 1–6
- Härmä, A., Jakka, J., Tikander, M., Karjalainen, M., Lokki, T., Hiipakka, J., et al. (2004). Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc.* 52, 618–639
- Hugill, A., and Amelides, P. (2016). "Audio only computer games—papa Sangre," In *Expanding the horizon of electroacoustic music analysis*. Editor S. Emmerson and L. Landy (Cambridge: Cambridge University Press), 355–375
- Klich, R. (2017). Amplifying sensory spaces: the in- and out-puts of headphone theatre. *Contemp. Theat. Rev.* 27, 366–378. doi:10.1080/10486801.2017.1343247
- Kochhar-Lindgren, K. (2002). Towards a communal body of art: the exquisite corpse and augusto boal's theatre. *Angelaki: J. Theoret. Humanities* 7, 217–226. doi:10.1080/09697250220142137
- Lyons, K., Gandy, M., and Starner, T. (2000). "Guided by voices: an audio augmented reality system," In Proceedings of international conference on auditory display, Boston, MA, April 2000 (Boston, MA: ICAD)
- Mariette, N. (2013). "Human factors research in audio augmented reality," In *Human factors in augmented reality environments*. Editors W. Huang, L. Alem, and M. A. Livingston (New York, NY: Springer New York), 11–32. doi:10.1007/978-1-4614-4205-9_2
- Mariette, N., and Katz, B. (2009). "Sounddelta—largescale, multi-user audio augmented reality," In EAA symposium on auralization, Espoo, Finland, June 15–17, 2009, 1–6
- McGinley, P. (2006). Eavesdropping, surveillance, and sound. *Perform. Arts J.* 28, 52–57. doi:10.1162/152028106775329642
- Moustakas, N., Floros, A., and Grigoriou, N. (2011). "Interactive audio realities: an augmented/mixed reality audio game prototype," AES 130th convention, London, UK, May 13–16, 2011, 2–8
- Peltola, M., Lokki, T., and Savioja, L. (2009). "Augmented reality audio for location-based games," AES 35th international conference on audio for games, London, UK, February 11–13, 2009, 6
- Playlines (2019). Consequences. [Dataset]
- Rovithis, E., and Floros, A. (2018). *Astrosonic: an educational audio gamification approach*. Corfu, Greece: Ionian University, 8
- Rovithis, E., Moustakas, N., Floros, A., and Vogklis, K. (2019). Audio legends: investigating sonic interaction in an augmented reality audio game. *MTI* 3, 73. doi:10.3390/mti3040073
- Saltz, D. Z. (2001). Live media: interactive technology and theatre. *Theat. Top.* 11, 107–130. doi:10.1353/tt.2001.0017
- Sawhney, N., and Schmandt, C. (2000). Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput. Hum. Interact.* 7, 353–383. doi:10.1145/355324.355327
- Schubert, T. W. (2003). The sense of presence in virtual environments: a three-component scale measuring spatial presence, involvement, and realism. *Z. Medienpsychol.* 15, 69–71. doi:10.1026//1617-6383.15.2.69
- Shin, D. (2019). How does immersion work in augmented reality games? A user-centric view of immersion and engagement. *Inf. Commun. Soc.* 22, 1212–1229. doi:10.1080/1369118X.2017.1411519
- Slater, M., and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoperators Virtual Environ.* 6, 603–616. doi:10.1162/pres.1997.6.6.603
- Tikander, M. (2009). Usability issues in listening to natural sounds with an augmented reality audio headset. *J. Audio Eng. Soc.* 57, 430–441
- Todd, F. (2019). Headphone theatre (performances & experiences): a practitioner's manipulation of proximity, aural attention and the resulting effect of paranoia. PhD Thesis. Essex: University of Essex
- Vazquez-Alvarez, Y., Oakley, L., and Brewster, S. A. (2012). Auditory display design for exploration in mobile audio-augmented reality. *Personal Ubiquitous Comput.* 16, 987–999. doi:10.1007/s00779-011-0459-0
- White, G. (2012). On immersive theatre. *Theat. Res. Int.* 37, 221–235. doi:10.1017/S0307883312000880
- Wilken, R. (2014). "Proximity and alienation: narratives of city, self, and other in the locative games of blast theory," In *The mobile story: narrative practices with locative technologies*. Editor J. Farman (New York: Routledge)
- Zimmermann, A., and Lorenz, A. (2008). LISTEN: a user-adaptive audio-augmented museum guide. *User Model. User-Adapted Interact.* 18, 389–416. doi:10.1007/s11257-008-9049-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Nagele, Bauer, Healey, Reiss, Cooke, Cowlshaw, Baume and Pike. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.